Where's the Beef

Does the Mack Method produce an undernourished range of possible outcomes?

Daniel Murphy, FCAS, MAAA
Trinostics LLC

CLRS 2009

In the GIRO Working Party's simulation analysis, actual unpaid losses exceeded the Mack Method's 99th percentile *over 10% of the time*

- *GIRO[1]:* The Mack Method "tends to understate the chance of extreme adverse outcomes, even in situations where the underlying assumptions are perfectly satisfied." (2007)
- *FSA[2]:* "Commonly used [stochastic] methods are inadequate to cover the full range of reserving variability." (2009)

1. http://www.actuaries.org.uk/__data/assets/pdf_file/0010/31303/BHPrize_Gibson.pdf
2. http://www.actuaries.org.uk/__data/assets/pdf_file/0009/146664/sm20090223.pdf

Agenda

1. Review GIRO WP simulation study
2. Analyze theory, visualize causes
3. Suggest improvements

$$Y = Xb + \sqrt{X}\sigma z$$

GIRO WP simulated 10,000 10x10 triangles …

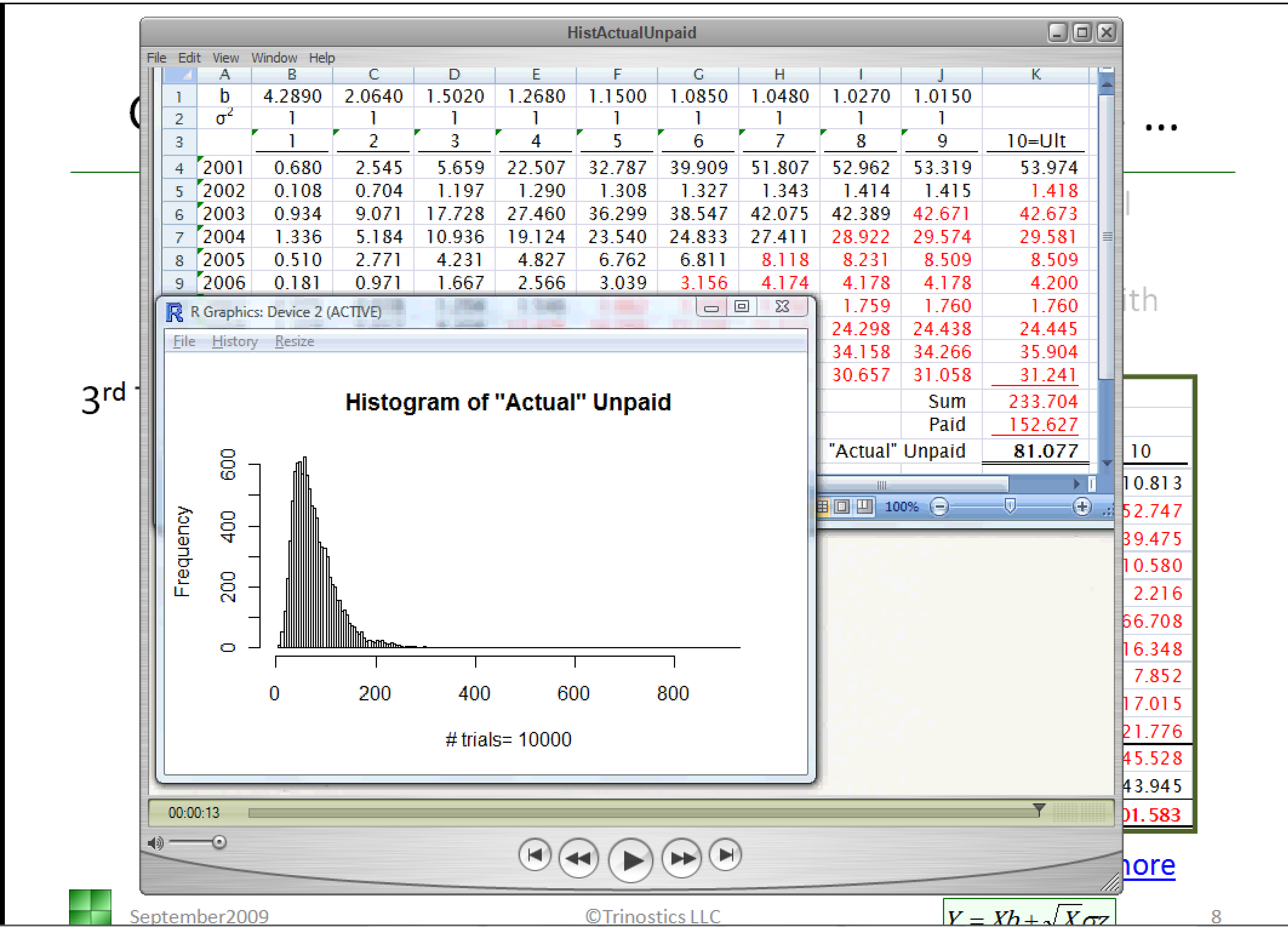
- Age 1 losses X_1: independent, random samples from a lognormal distribution with mean = 1, var = 1
- Age 2 losses X_2: randomly developed from a shifted lognormal with mean = $X_1 b_1$ and var = $X_1\sigma^2$. Similarly for age 3, 4, … , 10

1st Trial

b	4.2890	2.0640	1.5020	1.2680	1.1500	1.0850	1.0480	1.0270	1.0150	
σ^2	1	1	1	1	1	1	1	1	1	
AY \ age	1	2	3	4	5	6	7	8	9	10
2001	$0.420	$2.873	$7.175	$14.295	$16.676	$18.732	$36.983	$38.045	$38.102	$38.118
2002	0.824	3.742	8.875	12.354	14.176	14.680	15.044	15.080	15.111	15.347
2003	0.353	1.118	2.058	2.961	3.795	4.070	4.284	4.360	4.472	4.517
2004	2.669	8.403	12.937	18.463	24.133	24.811	25.665	25.725	25.729	25.873
2005	0.930	5.056	11.421	13.749	15.209	21.361	21.592	21.618	21.862	22.754
2006	0.357	1.382	2.485	3.002	3.135	6.299	6.455	6.523	6.550	6.565
2007	1.061	4.392	8.382	11.093	13.844	14.495	14.519	14.574	14.964	14.965
2008	1.308	6.626	10.563	14.934	20.307	20.861	26.047	26.345	26.398	26.406
2009	1.142	5.685	10.377	18.663	22.144	25.626	26.687	27.036	27.248	27.250
2010	1.639	7.667	16.534	27.154	37.583	40.970	64.295	64.772	65.366	65.411
									Sum	247.205
									Paid	136.731
									"Actual" Unpaid	**$110.474**

$Y = Xb + \sqrt{X}\sigma z$

Distribution of 10000 trials of actual unpaid claims



	A	B	C	D	E	F	G	H	I	J	K
1	b	4.2890	2.0640	1.5020	1.2680	1.1500	1.0850	1.0480	1.0270	1.0150	
2	σ^2	1	1	1	1	1	1	1	1	1	
3		1	2	3	4	5	6	7	8	9	10=Ult
4	2001	0.680	2.545	5.659	22.507	32.787	39.909	51.807	52.962	53.319	53.974
5	2002	0.108	0.704	1.197	1.290	1.308	1.327	1.343	1.414	1.415	1.418
6	2003	0.934	9.071	17.728	27.460	36.299	38.547	42.075	42.389	42.671	42.673
7	2004	1.336	5.184	10.936	19.124	23.540	24.833	27.411	28.922	29.574	29.581
8	2005	0.510	2.771	4.231	4.827	6.762	6.811	8.118	8.231	8.509	8.509
9	2006	0.181	0.971	1.667	2.566	3.039	3.156	4.174	4.178	4.178	4.200
									1.759	1.760	1.760
									24.298	24.438	24.445
									34.158	34.266	35.904
									30.657	31.058	31.241
										Sum	233.704
										Paid	152.627
									"Actual" Unpaid		81.077

… then ran 10,000 Chain Ladder, Mack Methods

1st Trial

Accident Year	Current Diagonal	Estimated VW ata	LDF	Estimated Ultimate	Estimated Unpaid	*Simulated Actual Unpaid*
2001	$ 38.118		1.0000	$ 38.118	$ -	$ -
2002	15.111	1.0004	1.0004	15.118	0.006	*0.235*
2003	4.360	1.002	1.002	4.369	0.009	*0.157*
2004	25.665	1.021	1.023	26.255	0.590	*0.208*
2005	21.361	1.316	1.346	28.757	7.396	*1.392*
2006	3.135	1.131	1.522	4.772	1.637	*3.430*
2007	11.093	1.190	1.811	20.089	8.996	*3.872*
2008	10.563	1.423	2.578	27.227	16.664	*15.844*
2009	5.685	1.902	4.903	27.874	22.189	*21.565*
2010	1.639	4.334	21.249	34.827	33.188	*63.772*
Sum	$ 136.731			$ 227.406	**$ 90.675**	*$ 110.474*

	Mack se	**39.085**	
	50%	90%	99%
Lognormal percentile	$83.269	$141.333	$217.550
Percentile sufficient this trial?	NO	YES	YES

$Y = Xb + \sqrt{X}\sigma z$

Insufficiency over all trials was … interesting!

Trial	Percentile			*Simulated Actual*	Percentile sufficient?		
	50%	90%	99%		50%	90%	99%
1	$ 83.269	$ 141.333	$ 217.550	*$ 110.474*	NO	YES	YES
2	78.329	120.532	171.279	*140.618*	NO	NO	YES
3	59.184	76.377	94.028	*101.583*	NO	NO	NO
…	…	…	…	…	…	…	…
9999	134.400	177.211	222.018	*40.433*	YES	YES	YES
10000	47.780	79.419	120.181	*41.434*	YES	YES	YES

	50%	90%	99%
Percent insufficient	***59.0%***	***25.6%***	***10.2%***
WP Table B-3	58.30%	24.55%	10.1%

$$Y = Xb + \sqrt{X}\sigma z$$

How can the Mack 99% VAR be so far off?

1. Is the *mean* of the distribution too low?
2. Is the *variance (MSE)* of the distribution too low?
3. Is the lognormal the wrong distribution to use?
4. Something else?

Is the chain ladder mean unpaid loss too low on average?

Accident Year	Theoretical mean unpaid	mean of actual (simulated) unpaid	% difference	mean of predicted unpaid	% difference
2001	$ -	$ -		-	
2002	0.340	0.351	3.4%	0.427	21.6%
2003	0.935	1.056	12.9%	1.021	-3.3%
2004	1.945	1.890	-2.8%	2.134	12.9%
2005	3.593	3.567	-0.7%	3.598	0.9%
2006	6.122	6.216	1.5%	6.166	-0.8%
2007	9.685	9.652	-0.3%	9.610	-0.4%
2008	14.129	14.155	0.2%	14.590	3.1%
2009	18.693	18.839	0.8%	18.981	0.8%
2010	21.982	22.284	1.4%	22.185	-0.4%
Sum	$77.422	$ 78.011	0.8%	$ 78.711	0.9%

- Chain ladder point estimate looks reasonably close to the actual mean value
- Not too low on average

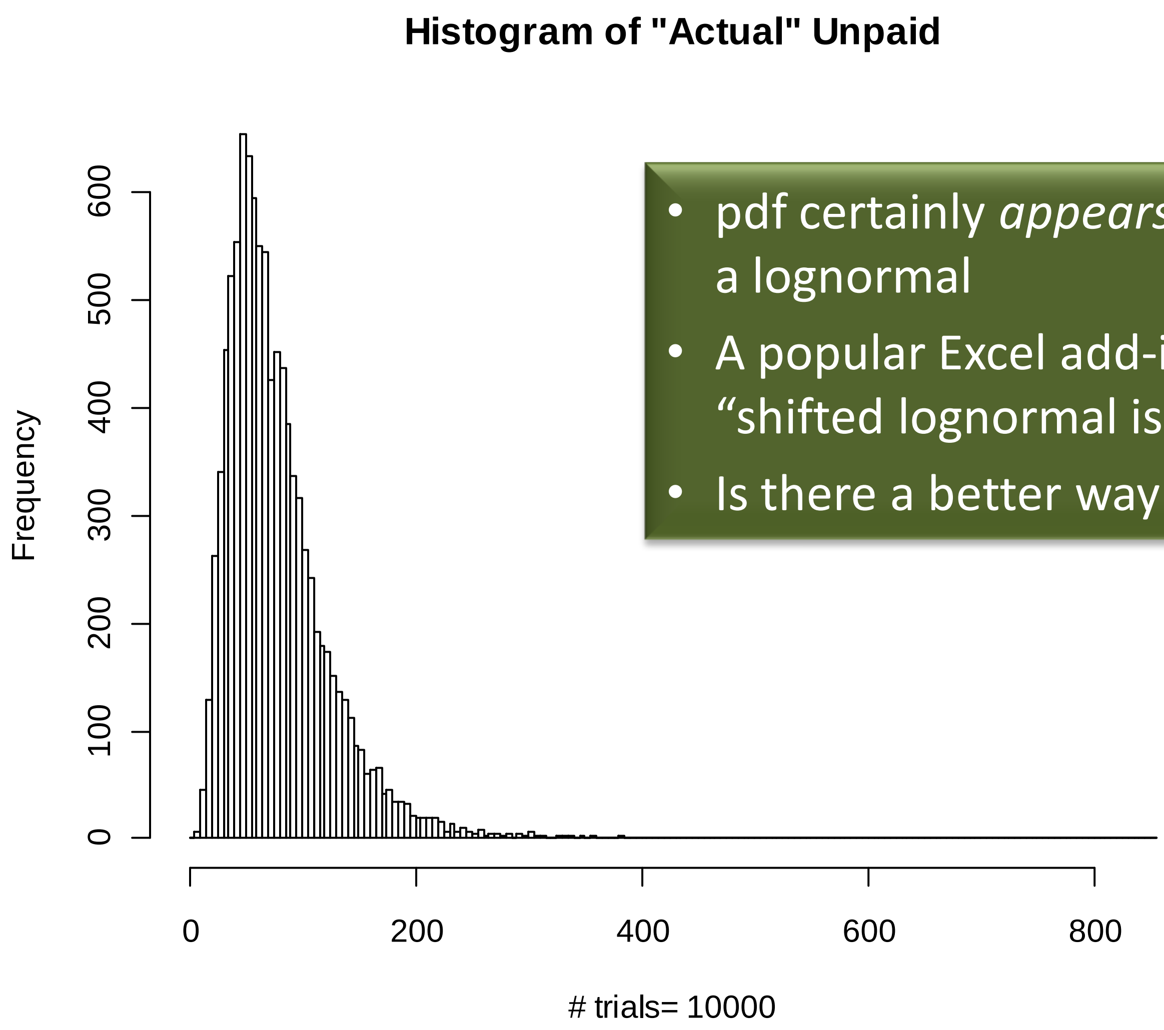
$Y = Xb + \sqrt{X}\sigma z$

Is the Mack Method standard error of unpaid loss too low on average?

Accident Year	Theoretical s.e. of unpaid	s.e. of actual (simulated) unpaid		s.e. of predicted unpaid	% difference
2001	$ -	$ -		-	
2002	NA	2.637		8.589	225.8%
2003	NA	6.200		9.149	47.6%
2004	NA	6.954		23.875	243.3%
2005	NA	8.198		9.761	19.1%
2006	NA	12.553		12.643	0.7%
2007	NA	14.163		13.348	-5.8%
2008	NA	18.893		41.590	120.1%
2009	NA	23.853		29.272	22.7%
2010	NA	27.333		24.980	-8.6%
Sum	NA	$ 46.757		$ 117.950	152.3%

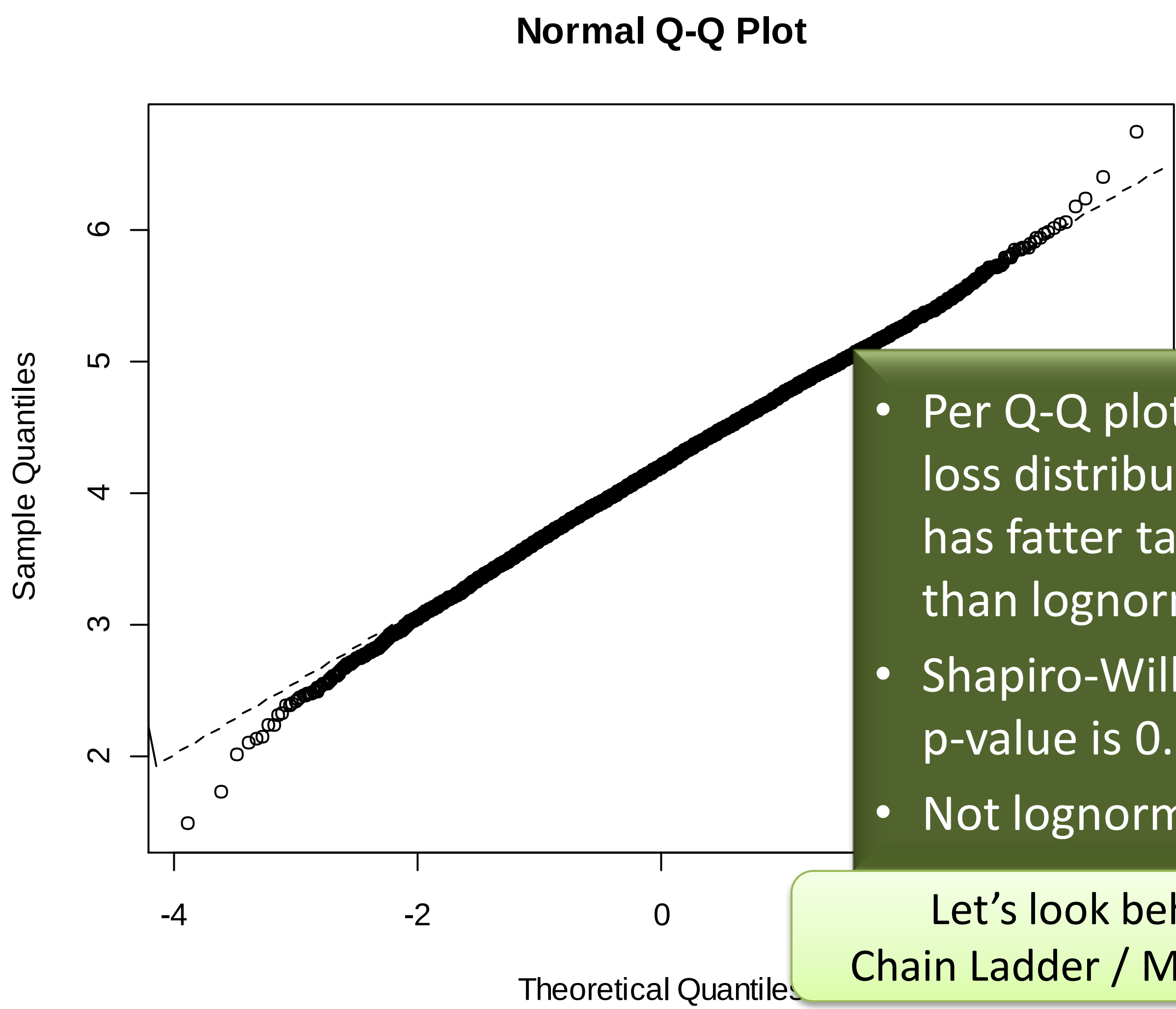
- Predicted variability does not appear too low on average

Do Algorithm A unpaid losses follow the lognormal distribution?



- pdf certainly *appears* to resemble a lognormal
- A popular Excel add-in says "shifted lognormal is best"
- Is there a better way to decide?

Do Algorithm A unpaid losses follow the lognormal distribution?



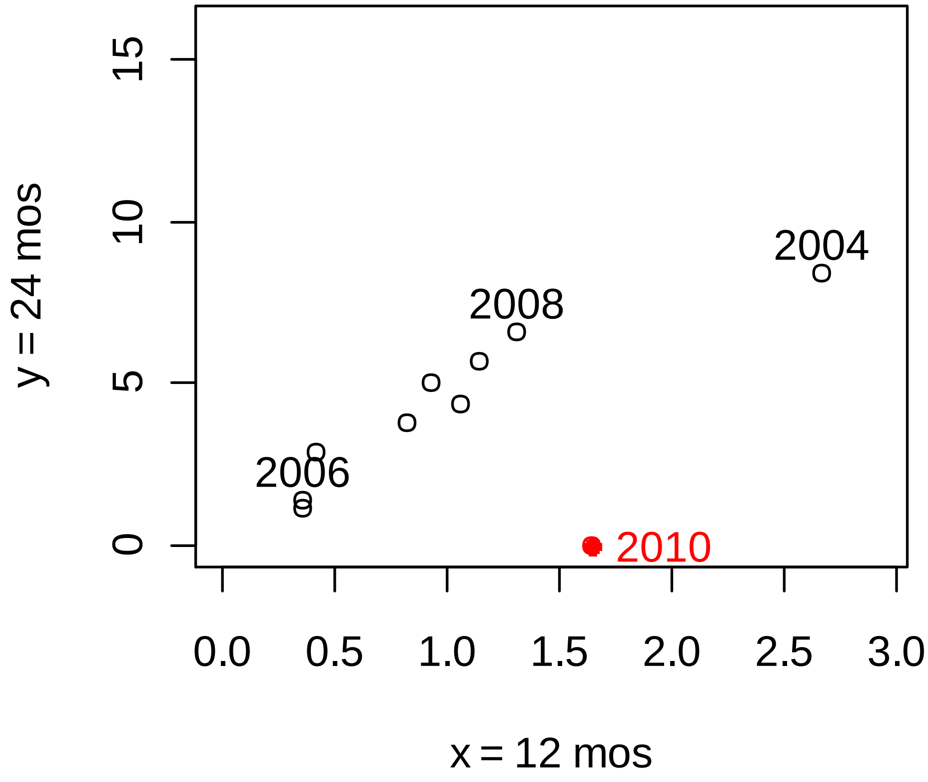
- Per Q-Q plot, unpaid loss distribution has fatter tails than lognormal
- Shapiro-Wilk test p-value is 0.006
- Not lognormal

Let's look behind the Chain Ladder / Mack Method

The CL method relates x and y values of loss

First Trial's age 12-24 development

AY	12	24
2001	$ 0.420	$ 2.873
2002	0.824	3.724
2003	0.353	1.118
2004	2.669	8.403
2005	0.930	5.056
2006	3.570	1.382
2007	1.061	4.392
2008	1.308	6.626
2009	1.142	5.685
2010	1.639	?



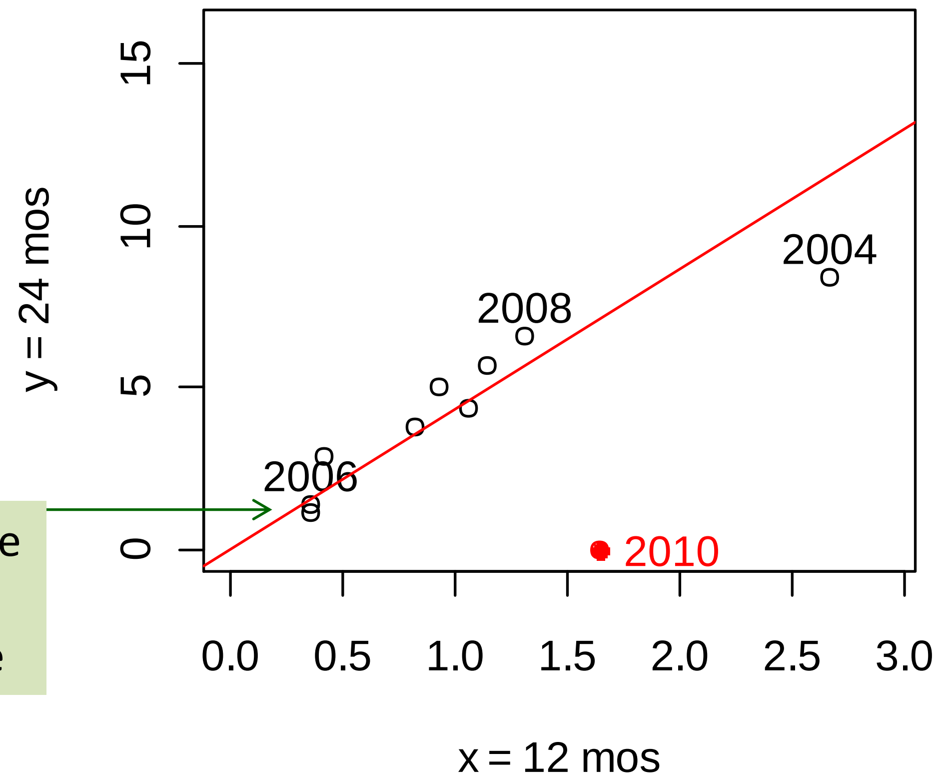
Volume weighted ATA is the slope of the line that represents that relationship

First Trial's age 12-24 development

AY	12	24
2001	$ 0.420	$ 2.873
2002	0.824	3.724
2003	0.353	1.118
2004	2.669	8.403
2005	0.930	5.056
2006	3.570	1.382
2007	1.061	4.392
2008	1.308	6.626
2009	1.142	5.685
2010	1.639	?

vol wtd ata
4.334

slope of line



Chain ladder projection of AY 2010 is the point on the estimated regression line

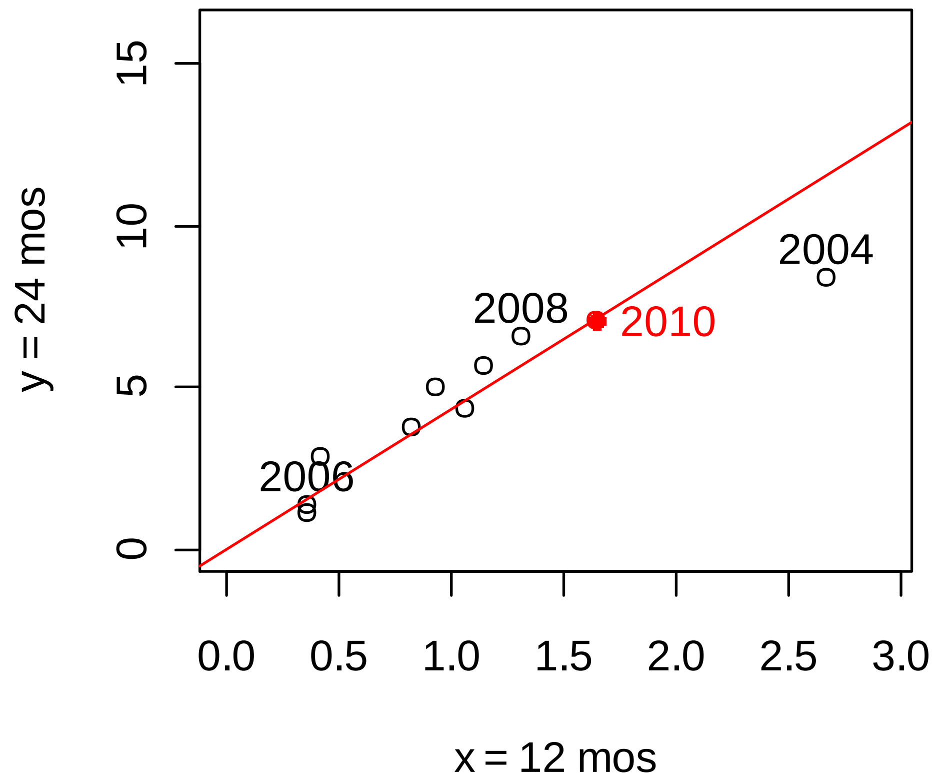
First Trial's age 12-24 development

Point Estimate

AY	12	24
2001	$ 0.420	$ 2.873
2002	0.824	3.724
2003	0.353	1.118
2004	2.669	8.403
2005	0.930	5.056
2006	3.570	1.382
2007	1.061	4.392
2008	1.308	6.626
2009	1.142	5.685
2010	1.639	7.103

vol wtd ata 4.334

$y = 4.334x$



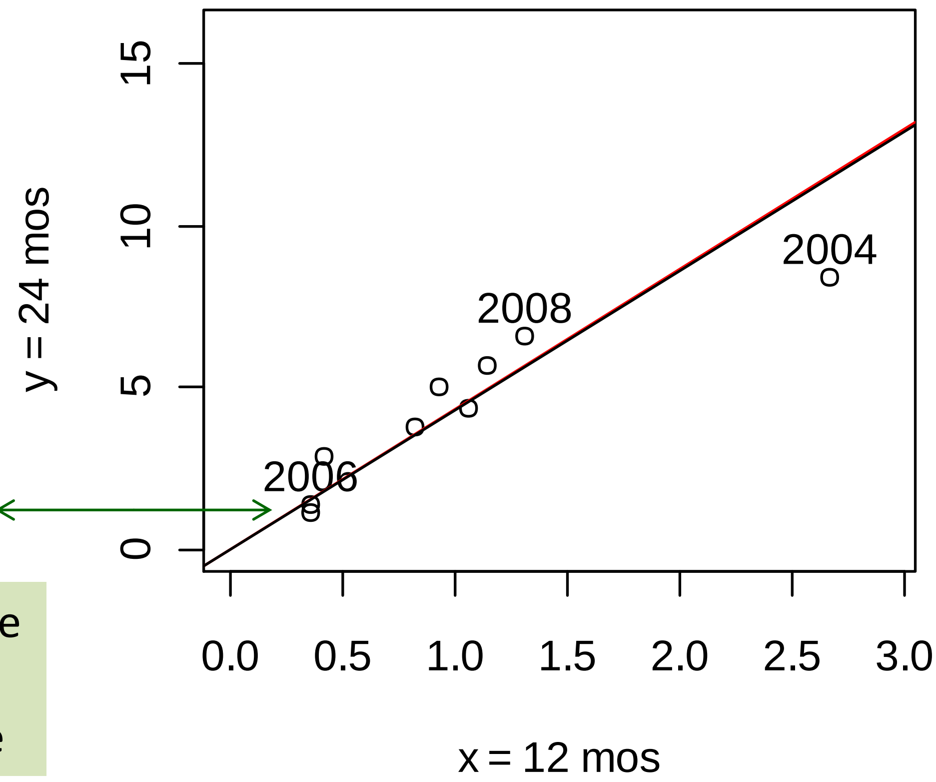
Blind luck that Trial 1's volume weighted ATA was so close to true slope

First Trial's age 12-24 development

AY	12	24
2001	$ 0.420	$ 2.873
2002	0.824	3.724
2003	0.353	1.118
2004	2.669	8.403
2005	0.930	5.056
2006	3.570	1.382
2007	1.061	4.392
2008	1.308	6.626
2009	1.142	5.685
2010	1.639	?

vol wtd ata 4.334 ("true" value = 4.289)

slope of line



Parameter risk reflects the fact that the slope was estimated from the data

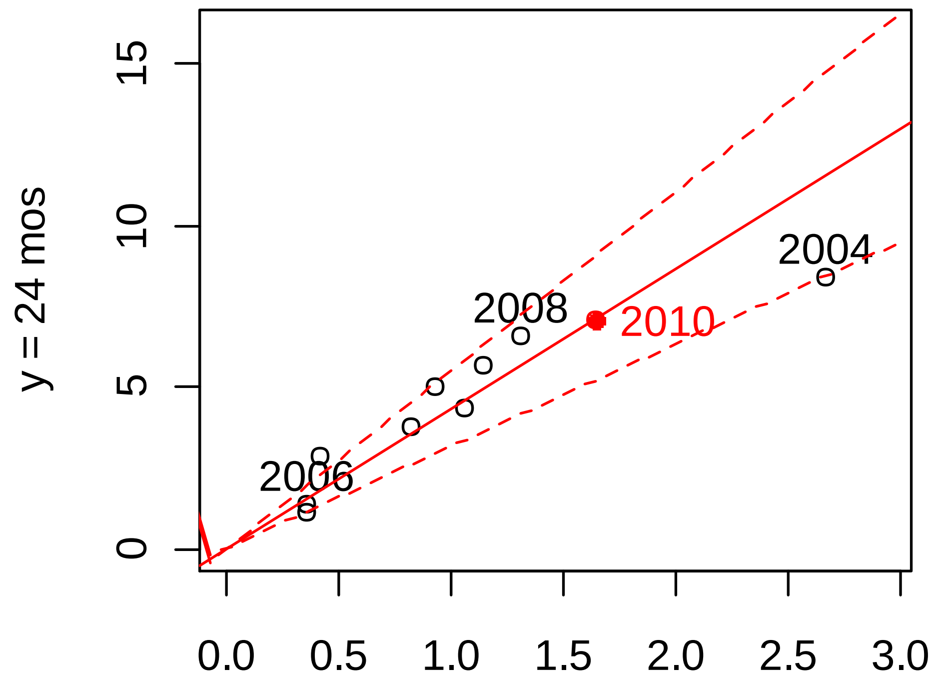
First Trial's age 12-24 development

AY	12	24
2001	$ 0.420	$ 2.873
2002	0.824	3.724
2003	0.353	1.118
2004	2.669	8.403
2005	0.930	5.056
2006	3.570	1.382
2007	1.061	4.392
2008	1.308	6.626
2009	1.142	5.685
2010	1.639	7.103

vol wtd ata 4.334

Parameter Risk

$y = (4.334 \pm 2*.359)x$



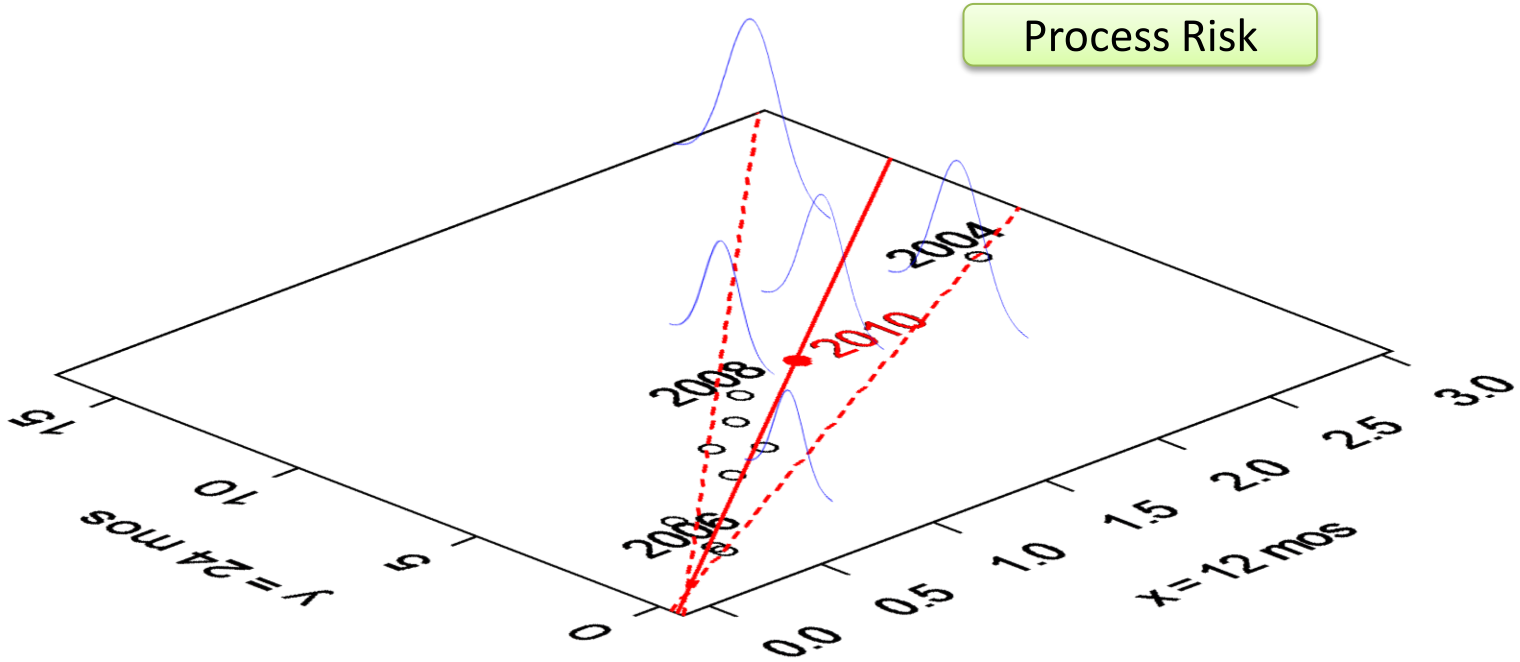
The dashed lines define a two-standard-error region within which the true line may fall

$Y = Xb + \sqrt{X}\sigma z$

Wherever the mean line truly is, losses will vary "noisily" around that expected value

First Trial's age 12-24 development

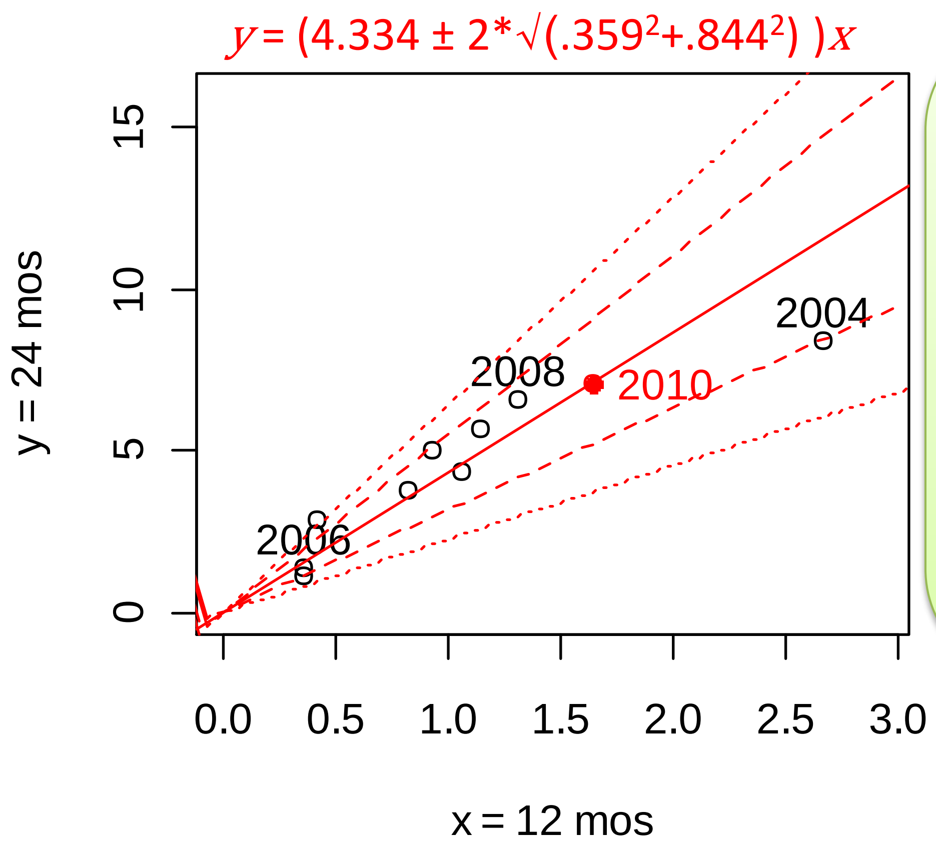
Process Risk



Total risk reflects both parameter & process risk

First Trial's age 12-24 development

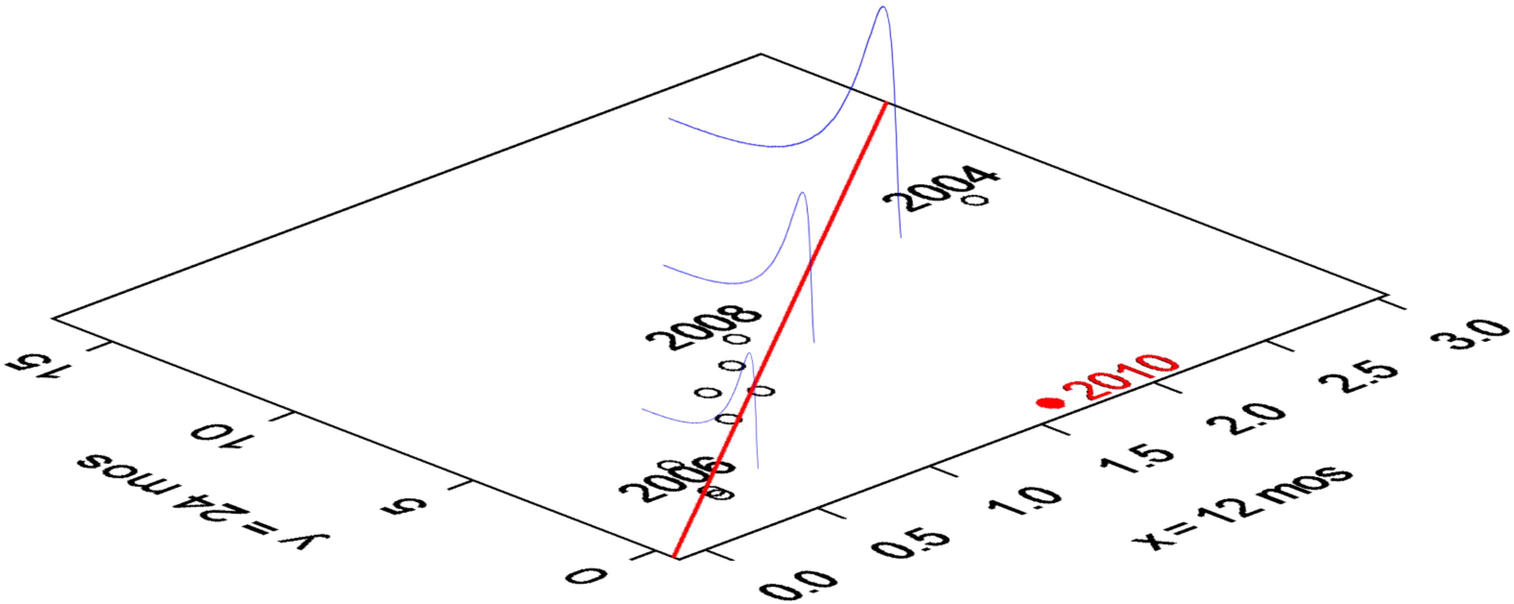
Total Risk



- The dotted lines define a two-standard-error region within which the next possible outcome may fall
- The statistics – 4.334, 0.359, 0.844 – can be derived from Excel's LINEST output

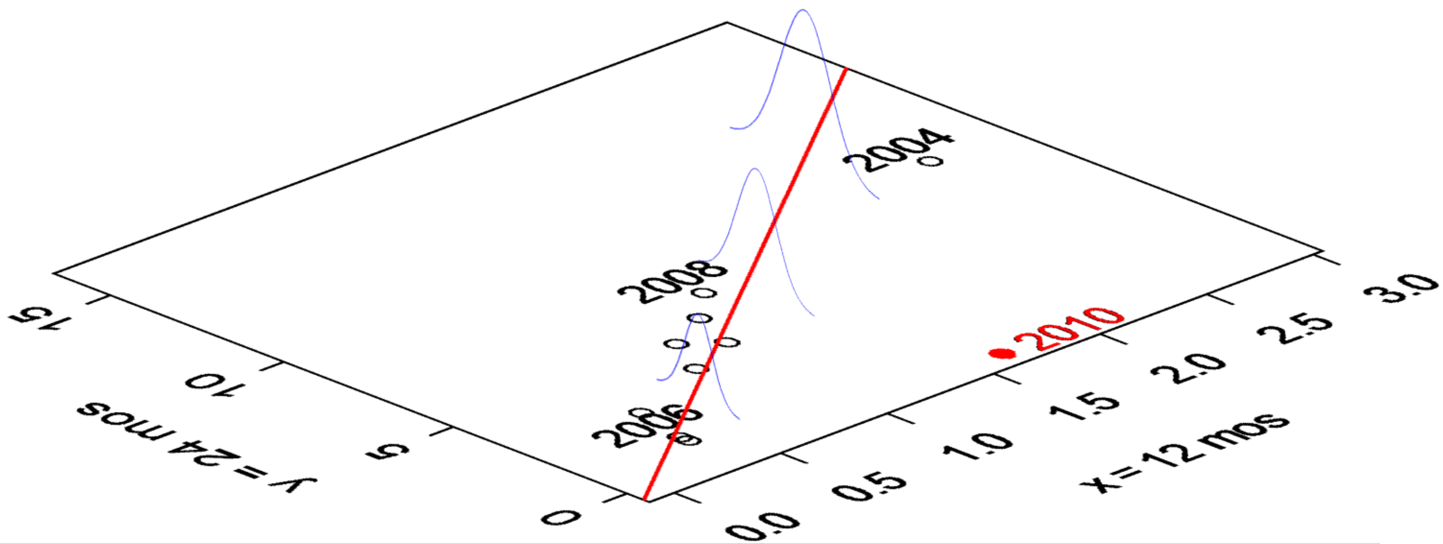
But Algorithm A development is not symmetric

First Trial's age 12-24 development



Algorithm A process risk is skewed
by virtue of lognormal assumption

If development were symmetric, 99% VAR insufficiency would drop from 10% to 2%



	Percent Insufficient		
	50%	90%	99%
Base case	58.49%	24.77%	10.18%
Link ratios ~ *Normal*	52.7%	13.7%	2.2%

We can't change reality. What can we do?

We should reflect uncertainty of spread parameter σ

Formulas for Development over a Single Period

Model: $Y = Xb + \sqrt{X}\sigma z$ for some unknown σ

Point estimate: $\hat{Y} = X\hat{b}$ where $\hat{b} = \Sigma y / \Sigma x$

Variance of link ratio: $Var(\hat{b}) = \sigma^2 / \Sigma x$

Variance of point estimate: $Var(\hat{Y}) = X^2\sigma^2 / \Sigma x$

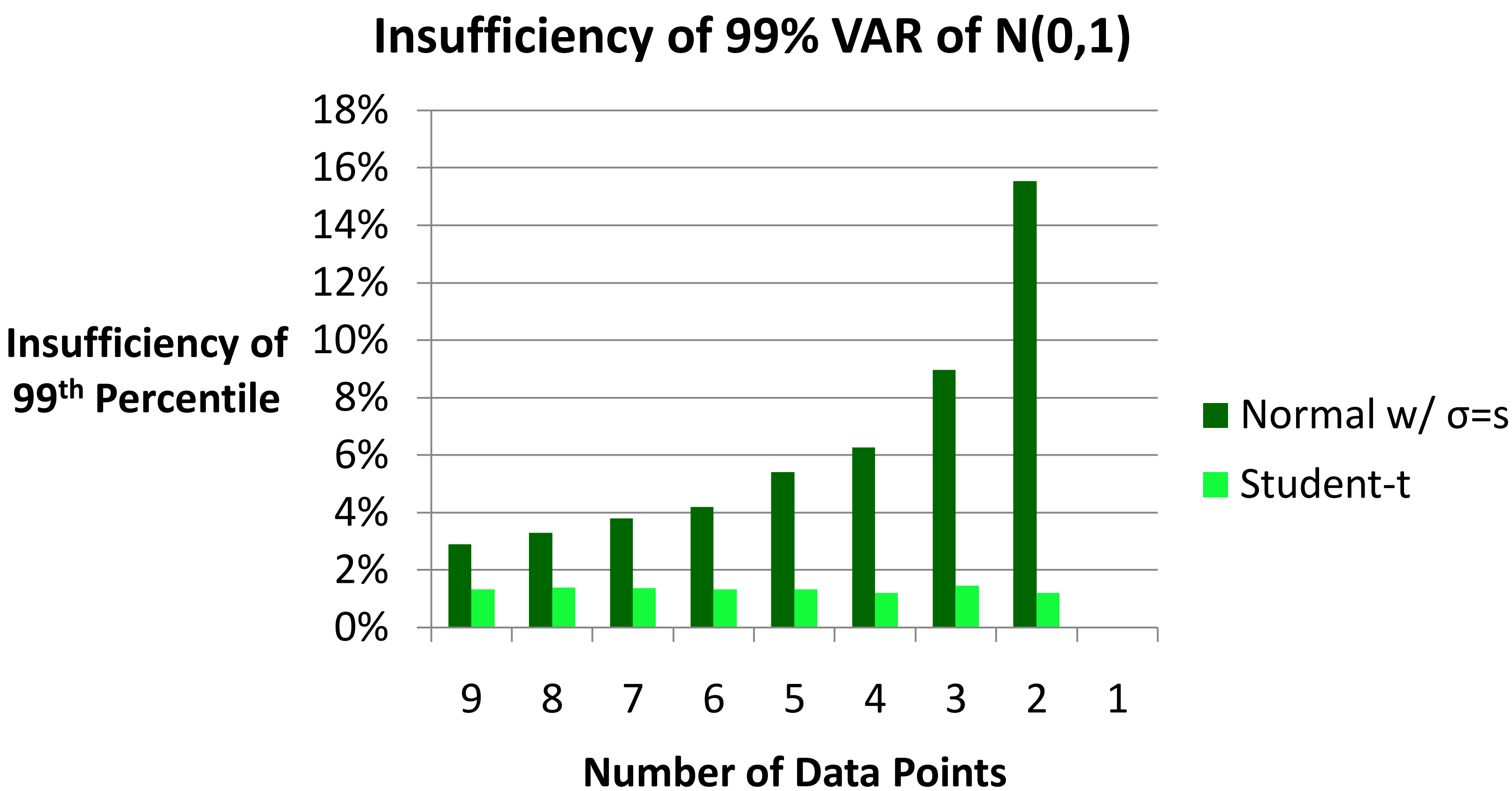
Variance of prediction: $Var(pred(Y)) = X^2\sigma^2 / \Sigma x + X\sigma^2$

GIRO

b	4.2890	2.0640	1.5020	1.2680	1.1500	1.0850	1.0480	1.0270	1.0150
σ^2	1	1	1	1	1	1	1	1	1

- Mack's formulas substitute the sample standard deviation s for σ in the above
- Practice understates estimated variability (Statistics 101)
- *student-t* distribution reflects spread uncertainty

Assuming spread is certain can significantly understate VAR, especially at mature ages



$Y = Xb + \sqrt{X}\sigma z$

We can simulate development at each age rather than fit a distribution at the end

Simulation can effectively and accurately replicate analytic results, including *student-t* based formulas

Distributions at each age can be chained together to develop the distribution of estimated unpaid claims

See for instance Gelman, *Data Analysis Using Regression and Multilevel/Hierarchical Models*

But, of all the alternative scenarios investigated, the most important discovery was …

With more accident years, insufficiency can be reduced significantly, even when

- facing a stacked deck (skewed development)
- spread is assumed known ($\sigma = s$)
- lognormal fit at the end

	Percent Insufficient		
	50%	90%	99%
0. Base case	58.49%	24.77%	10.18%
1. 3-term parameter risk formula	58.51%	24.76%	10.17%
2. Link ratios $\sim$ *Normal*	52.7%	13.7%	2.2%
3. Number of AY rows = 100	**48.3%**	**11.4%**	**2.7%**

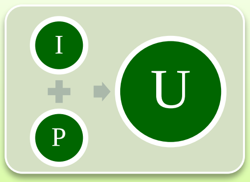
Conclusion

Reflect spread estimate uncertainty

t

- Particularly with small triangles (< 40 AYs)
- Adjust for degrees of freedom: t, chi-square distribution

Use analytic-equivalent simulation techniques

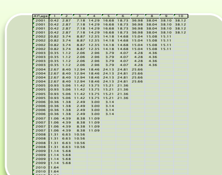


- Simulate hypothesized development
- Especially useful for complex models, even chain ladder

Test for skewed development

- Are residuals normally distributed?
- If not, try log transformation, GLMs, …

Analyze triangles at detailed levels



- More valuable information, more accurate projections
- By policy, claim, accident month, region

$Y = Xb + \sqrt{X}\sigma z$

Important advances in heart patient treatment have been made in the last ten years through more reliance on hard data and technical analysis and less reliance on expert opinion.

Dr. Raymond Stephens,
John Muir Medical Center Neurology,
Heart of Gold, 2009

Trinostics LLC is in the business of collaboration and education in the design and construction of valuable, transparent actuarial models

Daniel Murphy, FCAS, MAAA
dmurphy@trinostics.com